

## Original Article

# Whole genome association analysis shows that ACE is a risk factor for Alzheimer's disease and fails to replicate most candidates from Meta-analysis

Jennifer Webster<sup>1,6</sup>, Eric M. Reiman<sup>1,2,3,6</sup>, Victoria L. Zismann<sup>1,6</sup>, Keta D. Joshipura<sup>1,6</sup>, John V. Pearson<sup>1,6</sup>, Diane Hu-Lince<sup>1,6</sup>, Matthew J. Huentelman<sup>1,6</sup>, David W. Craig<sup>1,6</sup>, Keith D. Coon<sup>1,6,7</sup>, Thomas Beach<sup>6,8</sup>, Kristen C. Rohrer<sup>5</sup>, Alice S. Zhao<sup>5</sup>, Doris Leung<sup>5</sup>, Leslie Bryden<sup>5</sup>, Lauren Marlowe<sup>5</sup>, Mona Kaleem<sup>5</sup>, Diego Mastroeni<sup>8</sup>, Andrew Grover<sup>6,8</sup>, Joseph Rogers<sup>6,8</sup>, Reinhard Heun<sup>9</sup>, Frank Jessen<sup>9</sup>, Heike Kölsch<sup>9</sup>, Christopher B. Heward<sup>10</sup>, Rivka Ravid<sup>11</sup>, Michael L. Hutton<sup>12</sup>, Stacey Melquist<sup>12</sup>, Ron C. Petersen<sup>13</sup>, Richard J. Caselli<sup>6,14</sup>, Andreas Papassotiropoulos<sup>1,16</sup>, Dietrich A. Stephan<sup>1,2,6</sup>, John Hardy<sup>5,17</sup>, Amanda Myers<sup>4,5</sup>

<sup>1</sup>Neurogenomics Division, Translational Genomics Research Institute (TGen), Phoenix, AZ85004, USA; <sup>2</sup>Banner Alzheimer's Institute, Phoenix, AZ85006, USA; <sup>3</sup>Department of Psychiatry, University of Arizona, Tucson, AZ85724, USA; <sup>4</sup>Department of Psychiatry and Behavioral Sciences, University of Miami, Miller School of Medicine, Miami, FL33136, USA; <sup>5</sup>Laboratory of Neurogenetics, National Institute on Aging, National Institutes of Health, Bethesda, MD20892, USA; <sup>6</sup>Arizona Alzheimer's Consortium, Phoenix AZ85006, USA; <sup>7</sup>Division of Thoracic Oncology Research, St. Joseph's Hospital and Medical Center, Phoenix, AZ85013, USA; <sup>8</sup>Sun Health Research Institute, Sun City, AZ85351, USA; <sup>9</sup>Department of Psychiatry, University of Bonn, Sigmund-Freud-Strasse 25, 53105 Bonn, Germany; <sup>10</sup>Kronos Science Laboratory, Phoenix, AZ85016, USA; <sup>11</sup>Netherlands Institute for Neurosciences, Dutch Royal Academy of Arts and Sciences, Meibergdreef 47 AB Amsterdam, The Netherlands; <sup>12</sup>Department of Neuroscience, Mayo Clinic, Jacksonville, FL32224, USA; <sup>13</sup>Department of Neurology, Mayo Clinic, Rochester, MN55905, USA; <sup>14</sup>Department of Neurology, Mayo Clinic, Scottsdale, AZ85259, USA; <sup>15</sup>Department of Psychology, Arizona State University, Tempe, AZ85281, USA; <sup>16</sup>Division of Molecular Psychology and Life Sciences Training Facility, Biozentrum, University of Basel, Switzerland; <sup>17</sup>Reta Lila Weston Laboratories, Department of Molecular Neuroscience, Institute of Neurology, Queen Square, London WC1N3BG, England.

Received June 17, 2009; accepted July 20, 2009; available online July 26, 2009

**Abstract:** For late onset Alzheimer's disease (LOAD), the only confirmed, genetic association is with the apolipoprotein E (APOE) locus on chromosome 19. Meta-analysis is often employed to sort the true associations from the false positives. LOAD research has the advantage of a continuously updated meta-analysis of candidate gene association studies in the web-based AlzGene database. The top 30 AlzGene loci on May 1<sup>st</sup>, 2007 were investigated in our whole genome association data set consisting of 1411 LOAD cases and neuropathologically verified controls genotyped at 312,316 SNPs using the Affymetrix 500K Mapping Platform. Of the 30 "top AlzGenes", 32 SNPs in 24 genes had odds ratios (OR) whose 95% confidence intervals that did not include 1. Of these 32 SNPs, six were part of the Affymetrix 500K Mapping panel and another ten had proxies on the Affymetrix array that had >80% power to detect an association with  $\alpha=0.001$ . Two of these 16 SNPs showed significant association with LOAD in our sample series. One was rs4420638 at the APOE locus (uncorrected p-value=4.58E-37) and the other was rs4293, located in the angiotensin converting enzyme (ACE) locus (uncorrected p-value=0.014). Since this result was nominally significant, but did not survive multiple testing correction for 16 independent tests, this association at rs4293 was verified in a geographically distinct German cohort (p-value=0.03). We present the results of our ACE replication along with a discussion of the statistical limitations of multiple test corrections in whole genome studies.

**Key words:** Late-onset Alzheimer disease, single nucleotide polymorphism, genome-wide association study, meta-analysis, ACE

## Introduction

Until recently, most genetic analyses of complex traits involved candidate gene analysis.

Frequently, in such analyses, an initial positive report of a genetic association is followed by a mixture of positive and negative reports, leading to an unclear outcome after many studies. After such studies, it is often not clear whether there is an association, albeit weaker than that reported in the original study, or whether there is simply no association. For late-onset Alzheimer's disease (LOAD) (OMIM #104300), the only consistently replicated association locus is apolipoprotein E (*APOE*) [1,2]. Meta-analysis, with and without the inclusion of the original report, is commonly used to determine what are true and what are false positives. However, the original reports which are used to compile these meta-analyses may be subject to publication bias and it is not clear if positive reports of association are marginally more likely to be reported. Opinions are thus divided between those that believe that positive meta-analyses have considerable value ("glass half full") and those that believe these studies are too flawed by reporting bias to be genuinely useful ("glass half empty"). There has been little systematic analysis the utility of meta-analyses in any particular disorder in independent cohorts because the sample sizes and genotyping costs have been prohibitive.

In LOAD analysis we have the benefit of the AlzGene database [3], a continuously updated meta-analysis of all candidate gene association studies. This database allows us to assess large numbers of positive and negative meta-analyses simultaneously. With these resources we sought to assess whether the May 1<sup>st</sup> freeze of the top 30 Alzheimer meta-analyses (see **Table 1** for the list), 24 of which had 95% confidence intervals which did not include 1.0 could be translated into positive findings in our whole genome analyses using the 500K Affymetrix Mapping Array on 861 Alzheimer cases and 550 relevant controls. As previously reported, the most significant allele in this whole genome screen is rs4420638, which is located 14kb distal to the *APOE* locus [4]. We have also reported the association of a common haplotype of *GAB2* with LOAD in carriers of the *APOE*  $\epsilon$ 4 locus [5].

## Materials and methods

### Genotyping

A cohort of 1411 subjects including 643 brain donors who have been clinically diagnosed

and neuropathologically verified to have late-onset LOAD, 404 clinically and neuropathologically classified non-demented controls, 218 clinically diagnosed LOAD cases and 146 clinically characterized non-demented controls was utilized. The mean age of the AD cases was 82.8 $\pm$ 7.7 years and of controls was 79.7 $\pm$ 6.3 years. The 500K GeneChip (Affymetrix, Santa Clara, CA) was used to survey 502,267 SNPs in each subject as recently described [4, 5]. Genotypes were extracted using both SNIper-HD<sup>(6)</sup> and BRLMM (Affymetrix) software. 312,316 SNPs were analyzed after excluding those that were monomorphic, clustered poorly, had Hardy Weinberg equilibrium p-values less than 0.01, had minor allele frequencies less than 2%, or exhibited less than 98% concordance between the SNIper-HD and BRLMM calls. The software program STRUCTURE [7] was employed to test for underlying genetic stratification, using 5,000 randomly selected SNPs and including at least 100 SNPs per chromosome. The initial analysis yielded empirical evidence of three populations. Since fourteen subjects belonged to a population far removed from the rest of the study population, they were eliminated from further analyses. STRUCTURE then was used to demonstrate a comparable admixture of the two populations in the cases and controls. These data are at <http://www.tgen.org/research/index.cfm?pageid=1065> (see reference 5). ABI Taqman genotyping of rs4293 was carried out according to manufacturer's instructions, using the off-the-shelf kit, and read using an ABI 7000.

### Statistical Analysis

The PLINK analysis toolset (<http://pngu.mgh.harvard.edu/~purcell/plink/index.shtml>) was used for whole genome analysis. JMP was used to perform a logistic regression analysis of rs4293 vs. LOAD diagnosis with *APOE*  $\epsilon$ 4 gene dose as a covariate. LD mapping was performed by importing genotypes into the HaploView program version 3.32. Pair-wise LD values (as measured by  $D'$ ), reflect the likelihood that two genetic markers are inherited together. Power calculations were performed using the Genetic Power Calculator <http://pngu.mgh.harvard.edu/~purcell/gpc/> with a disease prevalence of 13%, and using AlzGene odds ratio (OR) as an estimator of relative risk [8].

### Follow up genotyping on an independent

**Table 1.** SNPs from Alzgene Freeze May 2008

AlzGene	AlzGene Polymorphism	Risk Allele	Freq. Cases	Freq. Controls	AlzGene OR (95% CI)	Affymetrix 500K SNP	r2	D'	Power (alpha=0.05)	Power (alpha=0.001)
<b>APOE</b>	<b>ε3/ε4</b>	<b>ε4</b>	<b>0.37</b>	<b>0.14</b>	<b>3.81 (3.38-4.29)</b>	<b>rs4420638</b>	<b>0.720</b>	<b>0.910</b>	<b>99.74%</b>	<b>92.92%</b>
CHRNA2	rs4845378 <sup>2</sup>	G	0.93	0.91	1.45 (1.06-1.96)					
<b>CH25H</b>	<b>rs13500</b>	<b>T</b>	<b>0.12</b>	<b>0.1</b>	<b>1.38 (1.01-1.88)</b>	<b>rs11594137</b>	<b>0.014</b>	<b>1.000</b>	<b>100.00%</b>	<b>100.00%</b>
<b>PGBD1</b>	<b>rs3800324</b>	<b>A</b>	<b>0.05</b>	<b>0.03</b>	<b>1.42 (1.13-1.80)</b>	<b>rs742107</b>	<b>1.000</b>	<b>1.000</b>	<b>99.94%</b>	<b>97.26%</b>
<b>LMNA</b>	<b>rs505058</b>	<b>C</b>	<b>0.08</b>	<b>0.06</b>	<b>1.35 (1.12-1.63)</b>	<b>rs2485668</b>	<b>1.000</b>	<b>1.000</b>	<b>100.00%</b>	<b>100.00%</b>
SOAT1	rs1044925	C	0.43	0.36	1.35 (1.13-1.60)	rs6666455	0.300	1.000	97.44%	73.19%
<b>MAPT</b>	<b>rs2471738</b>	<b>T</b>	<b>0.22</b>	<b>0.19</b>	<b>1.30 (1.01-1.67)</b>	<b>rs1078268</b>	<b>0.107</b>	<b>1.000</b>	<b>100.00%</b>	<b>100.00%</b>
<b>SORL1</b>	<b>rs1010159</b>	<b>C</b>	<b>0.36</b>	<b>0.34</b>	<b>1.14 (1.02-1.29)</b>	<b>rs1010159</b>	<b>1.000</b>	<b>1.000</b>	<b>100.00%</b>	<b>100.00%</b>
	<b>rs1699102</b>	<b>C</b>	<b>0.35</b>	<b>0.33</b>	<b>1.13 (1.02-1.25)</b>	<b>rs7116734</b>	<b>0.689</b>	<b>1.000</b>	<b>100.00%</b>	<b>99.99%</b>
	<b>rs2070045</b>	<b>G</b>	<b>0.26</b>	<b>0.23</b>	<b>1.26 (1.08-1.46)</b>	<b>rs11218347</b>	<b>0.015</b>	<b>1.000</b>	<b>99.28%</b>	<b>86.74%</b>
	<b>rs2276346</b>	<b>T</b>	<b>0.38</b>	<b>0.33</b>	<b>1.28 (1.07-1.54)</b>	<b>rs2276346</b>	<b>1.000</b>	<b>1.000</b>	<b>100.00%</b>	<b>99.89%</b>
	<b>rs2282649</b>	<b>T</b>	<b>0.3</b>	<b>0.28</b>	<b>1.16 (1.04-1.30)</b>	<b>rs2282649</b>	<b>1.000</b>	<b>1.000</b>	<b>100.00%</b>	<b>100.00%</b>
	<b>rs3824968</b>	<b>T</b>	<b>0.31</b>	<b>0.28</b>	<b>1.30 (1.07-1.58)</b>	<b>rs1629493</b>	<b>0.891</b>	<b>1.000</b>	<b>100.00%</b>	<b>100.00%</b>
	rs661057	T	0.59	0.56	1.14 (1.03-1.25)	rs610188	0.000	0.000	0.00%	0.00%
PCK1	rs8192708	A	0.13	0.11	1.29 (1.09-1.52)	rs1023049	0.066	0.782	77.33%	28.07%
CST3	rs1064039 <sup>2</sup>	A	0.21	0.19	1.16 (1.00-1.33)					
<b>ACE</b>	<b>rs1800764</b>	<b>T</b>	<b>0.59</b>	<b>0.54</b>	<b>1.27 (1.09-1.47)</b>	<b>rs4293</b>	<b>0.964</b>	<b>1.000</b>	<b>99.85%</b>	<b>94.85%</b>
	rs4291 <sup>2</sup>	T	0.65	0.62	1.22 (1.04-1.43)					
SORCS1	rs600879 <sup>2</sup>	min	0.11	0.09	1.24 (1.04-1.48)					
<b>TF</b>	<b>rs1049296</b>	<b>C2</b>	<b>0.19</b>	<b>0.17</b>	<b>1.24 (1.06-1.25)</b>	<b>rs1049296</b>	<b>1.000</b>	<b>1.000</b>	<b>100.00%</b>	<b>100.00%</b>
hCG2039140	rs1903908 <sup>2</sup>	T	0.15	0.12	1.23 (1.06-1.44)					
IDE <sup>1</sup>										
GALP	rs3745833	C	0.39	0.35	1.21 (1.10-1.33)	rs4801296	0.187	0.598	94.68%	61.20%
<b>CTSD</b>	<b>rs17571</b>	<b>T</b>	<b>0.09</b>	<b>0.08</b>	<b>1.20 (1.01-1.42)</b>	<b>rs17834326</b>	<b>0.786</b>	<b>1.000</b>	<b>100.00%</b>	<b>100.00%</b>
TNK1	rs1554948	T	0.55	0.5	1.19 (1.08-1.32)	rs7219773	0.967	1.000	83.86%	36.62%

ACE as AD Risk Factor

(Continued Table 1)

GWA_14q32.13	rs11622883	T	0.58	0.54	1.19 (1.08-1.30)	rs17091290	0.008	1.000	7.58%	0.25%
<b>IL1B</b>	rs1143634	T	0.28	0.26	1.18 (1.02-1.37)	rs3917365	0.013	1.000	97.75%	75.01%
LOC651924	rs6907175 <sup>2</sup>	G	0.49	0.45	1.16 (1.04-1.30)					
PON1 <sup>1</sup>										
GWA_7p15.2 <sup>1</sup>										
LOC439999	rs498055	G	0.52	0.48	1.15 (1.03-1.29)	rs526928	0.887	1.000	15.89%	0.97%
<b>DAPK1</b>	<b>rs4877365</b>	<b>G</b>	<b>0.74</b>	<b>0.69</b>	<b>1.25 (1.09-1.45)</b>	<b>rs4877365</b>	<b>1.000</b>	<b>1.000</b>	<b>100.00%</b>	<b>100.00%</b>
	rs4878104	C	0.66	0.62	1.15 (1.05-1.27)	rs4877365				
PRNP	rs1799990	A	0.69	0.66	1.14 (1.03-1.27)	rs7274758	0.048	1.000	16.73%	1.07%
MYH13 <sup>1</sup>										
HMGCS2 <sup>1</sup>										
<b>BDNF</b>	<b>rs6265</b>	<b>A</b>	<b>0.19</b>	<b>0.18</b>	<b>1.10 (1.01-1.19)</b>	<b>rs6265</b>	<b>1.000</b>	<b>1.000</b>	<b>100.00%</b>	<b>100.00%</b>
PSEN1 <sup>1</sup>										

<sup>1</sup>These genes had no polymorphisms with 95% confidence intervals that did not include 1. <sup>2</sup>These SNPs were not genotyped by the HapMap Consortium and therefore we were unable to determine whether there was a sufficiently powerful proxy on the Affymetrix 500K Mapping Array with which to test the association. **Bolded** SNPs (>80% power to detect  $\alpha=0.001$ ) were examined in our whole genome association data set.

*sample series*

We used a follow up sample of LOAD patients recruited from the Department of Psychiatry of the University of Bonn Patients. These were diagnosed according to DSM IV criteria, supported by clinical examination and detailed. Healthy controls from the general population were recruited with the support of the local Census Bureau and the regional Board of Data protection (Nordrhein-Westfalen, Germany). Neuropsychological testing, detailed structured interviews and clinical examination were performed. All patients and control subjects gave informed consent for participation in the study. The study protocol was approved by the Ethics Committee of the Faculty of Medicine of the University of Bonn. The LOAD patients (n=200) had a mean age of 76±6.5 years and a mean age at onset of 73±7.1 years and 72.5% were women and the controls (n=126) had a mean age of 73±6.4 years and 50.8 % were women. With the exception of their national origin, these subjects are similar to those we used in our whole genome study [5].

**Results**

Our first step was to determine the extent to which our study had 80% power to replicate

the reported associations with an alpha of 0.001 and allow for replication after multiple testing correction. This relied on two factors other than the size of our sample: the size of the reported effect and the extent to which a proxy SNP on the Affymetrix array is in linkage disequilibrium with the SNP used in the candidate gene study. These parameters are illustrated in **Table 1**. As these data show, the SNP on the array was the same as the candidate SNP in six cases (rs1010159, rs2276346 and rs2282649 in *SORL1*, rs1049296 in *TF*, rs4877365 in *DAPK1*, and rs6265 in *BDNF*) and was in sufficient LD with that SNP ( $r^2$  0.014 to 1.00,  $D'$  0.91 to 1.00, power to detect association at  $\alpha=0.001$  80% to 100.00%) in a further ten cases. The uncorrected contingency test p-values for these sixteen SNPs are listed in **Table 2**. In eight cases the Affymetrix platform does not capture the variability of the reported SNP with sufficient power for our study to be a useful proxy.

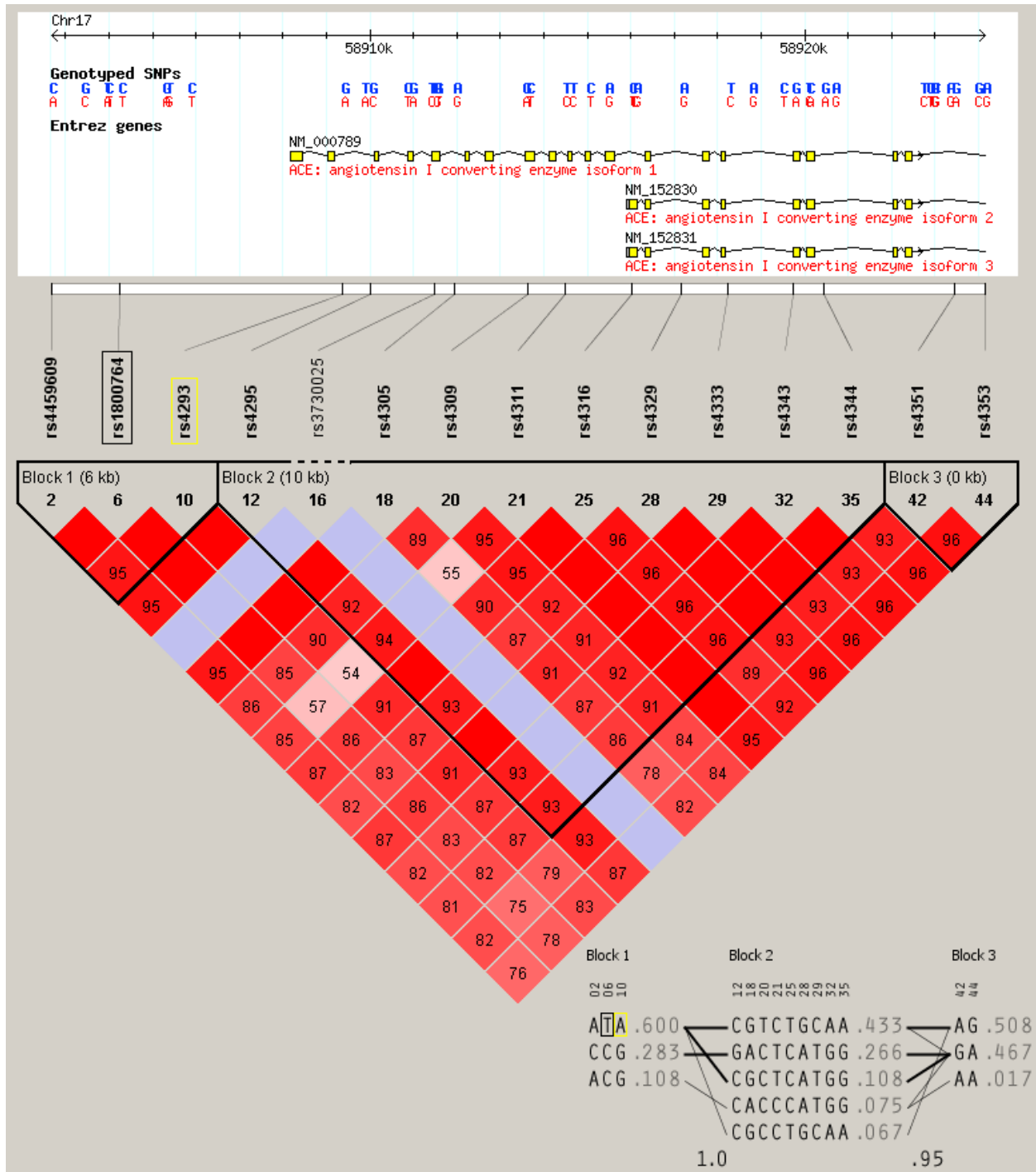
As we have previously reported, and unsurprisingly, this analysis confirms that *APOE* is a risk locus for disease, since the genetic association between the  $\epsilon 4$  allele is captured by rs4420638 ( $r^2=0.72$ ). Of more interest, is the fact that in only one of the other 15 cases

**Table 2.** Data from ALZgene-related SNPs in Whole Genome Analysis

AlzGene	Affymetrix 500K SNP	Risk Allele	Freq. Cases	Freq. Controls	P-value	OR (95% C.I.)
APOE	rs4420638	C	0.40	0.17	4.58E-37	3.2 (2.68-3.9)
CH25H	rs11594137	A	0.86	0.85	0.7159	
PGBD1	rs742107	C	0.061	0.050	0.2319	
LMNA	rs2485668	A	0.95	0.95	0.2833	
ACE	rs4293	A	0.57	0.52	0.014	1.22 (1.04-1.41)
TF	rs1049296	C	0.82	0.82	0.82	
DAPK1	rs4877365	C	0.71	0.70	0.76	
BDNF	rs6265	T	0.21	0.21	0.65	
MAPT	rs1078268	A	0.79	0.78	0.72	
SORL1	rs1010159	T	0.67	0.64	0.12	
SORL1	rs7116734	G	0.60	0.58	0.30	
SORL1	rs11218347	T	0.92	0.93	0.91	
SORL1	rs2276346	G	0.64	0.63	0.26	
SORL1	rs2282649	C	0.71	0.70	0.39	
SORL1	rs1629493	A	0.64	0.61	0.092	
CTSD	rs17834326	T	0.099	0.092	0.56	

Uncorrected, two-tailed contingency test p-values for the sixteen Affymetrix 500K SNPs that are good proxies for significant polymorphisms in the AlzGene database. The "Risk allele" is that allele reported to be associated with disease in previous studies or to be the Affymetrix surrogate for that allele.

# ACE as AD Risk Factor



**Figure 1.** Linkage disequilibrium structure of HapMap genotyped SNPs in the CEPH population at the distal end of ACE. The plot follows the standard HAPLOVIEW color scheme: white boxes,  $D' < 1$ ,  $LOD < 2$ ; shades of pink/red,  $D' < 1$ ,  $LOD \geq 2$ ; blue,  $D'=1$ ,  $LOD < 2$ ; bright red,  $D'=1$ ,  $LOD \geq 2$ . AlzGene SNP rs1800764 (black box) is part of the same haplotype block as Affymetrix SNP rs4293 (yellow box). The AD risk allele (T, in black box) associated with rs1800764 is transmitted with the AD risk allele (A, in yellow box) that was detected in rs4293 by this screen.

was the association confirmed, and this was at the *ACE* locus (also known as DCP1). The Affymetrix 500K SNP at the *ACE* locus that we tested, rs4293, was nominally significant (uncorrected contingency test p-value=0.014) in our study population.

According to the HapMap CEPH data set (9), rs4293 is part of the same haplotype block as rs1800764, the significant polymorphism identified in the AlzGene database (**Figure 1**). Additionally, the risk allele for rs1800764 (T) is part of the same haplotype as the risk allele we identified in rs4293 (A), confirming that the risk allele for rs4293 displayed the same direction of effect as the identified risk allele of rs1800764.

Given the nominal significance of the *ACE* replication, a second cohort was examined to determine the true significance of *ACE* in LOAD. A replication cohort of clinically assessed patient samples (see Methods) from Bonn, Germany was genotyped at rs4293 using an ABI Taqman (Applied Biosystems, Foster City, CA) custom genotyping assay. An association with LOAD was confirmed in this population at rs4293 with the same direction of effect observed in our whole genome association series (**Table 3**).

**Table 3.** Association of rs4293 with AD in the Bonn Alzheimer Series

	Genotypes (AA/AG/GG)	A Allele Frequency	G Allele Frequency
Cases	75/89/36	0.60	0.40
Controls	34/61/31	0.51	0.49
Total	109/150/67	0.56	0.44

Association of rs4293 with AD in a clinically characterized cohort from Bonn, Germany. The odds ratio for the G allele was 0.71 (95% confidence interval, 0.51-0.97). The 'p' value (two tailed t test) was 0.03.

A logistic regression analysis was used to test for an interaction between rs4293 and *APOE*  $\epsilon$ 4 gene dose in the determination of LOAD risk. The p-value of this interaction term was 0.3567, indicating that there is not a significant interaction between *ACE* and *APOE* in determining disease risk in this cohort.

## Discussion

These data are of interest for three reasons. First, they show that the *ACE* locus is indeed

an Alzheimer risk factor. Second, they illustrate that Bonferroni correction leads to false negative findings in the context of whole genome associations since this risk factor was missed in our initial analysis. Lastly, they illustrate that meta analyses lead to reports of positive associations which fail to replicate even when samples series of sufficient power and correct ethnicity are used, presumably because of bias in the reporting of data.

### *ACE as an Alzheimer Risk factor*

Kehoe and colleagues (1999) reported an association of an insertion/deletion polymorphism in intron 16 of *ACE* with LOAD [10] that has been shown to account for about half of *ACE* serum expression levels [11]. Subsequently, seven polymorphisms in *ACE* were examined, five of which are included in the AlzGene analysis. Of these five, two (rs4291 and rs1800764) were significant according to our criteria. Since rs4291 was not genotyped as a part of the HapMap initiative, we were unable to determine if it had a good proxy on the Affymetrix 500K Mapping Array. However, there is a good proxy for rs1800764 in rs4293, for which we have 95% power to detect the observed association with LOAD. A nominally significant association was detected in the whole genome association cohort (uncorrected two-tailed p-value=0.014) with the same direction of effect observed in the AlzGene meta-analysis. Since this association did not survive a Bonferroni multiple testing correction, we confirmed the association of rs4293 with LOAD in a second, geographically distinct population (two-tailed p-value=0.03). This association had the same direction of effect as our whole genome association series and the AlzGene analysis.

*ACE* is hypothesized to be a risk factor for atherosclerosis [12] and hypertension [13]. There is now epidemiological [14-16] and pathological [17-19] evidence pointing to vascular risk factors for LOAD. A genetic association with the allele of the *ACE* locus which is associated with risk for hypertension [13, 20], fits with this pattern and with the hypothesis that microhaemorrhages may play a part in Alzheimer pathogenesis [21].

### *Bonferroni correction and whole genome studies*

These data show, unsurprisingly, that Bonfer-

roni correction for multiple tests, leads to false negative findings in whole genome studies. The *ACE* polymorphism was the 4398th out of 312,316 in terms of the significance of the allelic 'p' values. Here we show that *ACE* is a risk factor for LOAD, however using a traditional whole genome association analysis approach, including a Bonferroni correction, we did not identify *ACE* [5]. A great deal of attention has been paid, with good reason, to limiting type I error in genome-wide scans [22, 23]. However, clearly and unsurprisingly the use of the Bonferroni correction greatly increases the type II error and decreased the power of our whole genome association dataset [24, 25]. This, and the recent success of the whole genome scans for diabetes [26-29] illustrates that a productive way forward for identifying loci which do not appear to stand out considerably from the mass of data, such as *APOE* does in this data set for LOAD [4], is to have large initial cohorts for whole genome analysis followed by large and independent replication cohorts in which replication of several thousand SNPs (at least 5000) is attempted, followed by third cohorts etc.

#### *The (Limited) Utility of Meta-Analyses*

Finally, this study suggests that meta-analyses have modest utility in distinguishing true from false positives in genetic associations. *ACE* was the 11<sup>th</sup> highest ranked by allelic effect size (<http://www.alzforum.org/res/com/gen/alzgene/methods.asp#topresults>). The most likely reason for the failure to replicate the other candidates from meta analysis is that even subtle publication biases are likely to be enough to lead to the false assignment of statistical significance to odds ratios in the range of 1.1 to 1.3. Even the most robust meta-analysis methods are limited by the literature available. Meta-analyses assume that the literature represents a random sampling of studies for each association marker. However, if negative findings are underreported in comparison to positive association findings, meta-analyses may yield false positives or inflated odds ratios [30, 31]. A recent meta-analysis of twenty-five associations across multiple complex disorders found statistically significant replication for eight of the reported associations, or 32% of reported associations [32], and concluded that publication bias was not responsible for the observed associations. However, that analysis did not attempt external

testing of these associations as we have done here and made the assumption that, following a positive report of association, associations reporting replication and failure to replicate were equally likely to be published. In contrast, we only identified sixteen SNPs from the AlzGene meta-analysis for which we had greater than eighty percent power to detect the reported association in our whole genome association dataset. Of these, only two (13%), *APOE*  $\epsilon$ 4 and rs4293 in *ACE*, were significantly associated with AD in our large case-control population. Clearly, one could argue that meta-analysis is not required for confirmation of the *APOE* association and this would reduce the predictive success rate for meta-analyses to be 1/15 (7%) in this disorder at least. While this may seem rather pessimistic, it suggests that systematic meta-analyses of the literature, such as AlzGene, are useful adjuncts to whole genome studies in the identification of risk alleles. It is notable that, subsequent to these analyses, *ACE* has moved to number two in the Alzgene list (September 1<sup>st</sup> 2009) and that others have more recently suggest *ACE* is an important disease risk factor, in part because of their analysis of whole genome data [33-35].

#### *Summary*

As data from other diseases have indicated, whole genome association studies have the potential to identify genetic risk variants which contribute to disease. The data we present here show that genetic variability in *ACE*, does, as others had suggested [10, 30] contribute to the risk of LOAD and indicate that, while previous meta-analyses cannot be unequivocal determinants of whether genetic variability at a certain locus contributes to risk, they are a useful step as hypothesis generating analyses which can then be tested in a whole genome study.

#### **Acknowledgements**

These studies were supported by Kronos Life Sciences Laboratories, the National Institute on Aging (Arizona Alzheimer's Disease Center P30 AG19610, RO1 AG023193, Mayo Clinic Alzheimer's Disease Center P50 AG16574 and Intramural Research Program), the National Alzheimer's Coordinating Center (U01 AG016976), and the state of Arizona. We thank our research volunteers and their families for their generous participation and Drs. Creighton



Phelps, Marcelle Morrison-Bogorad, Marilyn Miller, and Walter Kukull for their assistance in the acquisition of tissue samples and data, and directors, pathologists, and technologists from the following ADCs and brain banks: Lucia Sue (Sun Health Research Institute and Arizona Alzheimer's Disease Center); Ruth Seemann and Dan Brady (National Institute on Aging); Juan C. Troncoso and Olga Pletnikova (John Hopkins, P50 AG05146); Harry Vinters and Justine Pomakian (University of California, Los Angeles, P50 AG16570); Christine Hulette (The Kathleen Price Bryan Brain Bank, Duke University Medical Center, P50 AG05128, R01 NS39764, R01 MH60451, and GlaxoSmith-Kline); Dikran Horoupian, Ahmad Salehi (Stanford University, P30 AG17824); Jean Paul Vonsattel (New York Brain Bank, Taub Institute, Columbia University, P50 AG08702); E. Tessa Hedley-Whyte, Karlotta Fitch (Massachusetts General Hospital, P50 AG05134); Roger Albin, Lisa Bain, and Eszter Gombosi (University of Michigan, P50 AG08671); William Markesbery, Sonya Anderson (University of Kentucky, P50 AG05144); Dennis W. Dickson and Natalie Thomas (Mayo Clinic, Jacksonville, P50 AG16574 and P50 AG25711); Carol A. Miller, Jenny Tang, and Dimitri Diaz (University of Southern California, P50 AG05142); Dan McKeel, John C. Morris, Eugene Johnson, Jr., Virginia Buckles, and Deborah Carter (Washington University, St Louis, P50 AG 05681); Thomas Montine and Aimee Schantz (University of Washington, P50 AG05136); John Q Trojanowski, Virginia M Lee, Vivianna Van Deerlin, and Terry Schuck (University of Pennsylvania); Ann C. McKee and Carol Kubilus (Boston University, P30 AG13846); Bruce H. Wainer and Marla Gearing (Emory University, AG025688); Charles L. White, III, Roger Rosenberg, Marilyn Howell, and Joan Reisch (University of Texas, Southwestern Medical School, P30-AG12300); William Ellis and Mary Ann Jarvis, (University of California, Davis, P30 AG AG01542); David A. Bennett, Julie A. Schneider, Karen Skish, and Wayne T. Longman (Rush University Medical Center, P30 AG10161); and Deborah C. Mash, Margaret J Basile, and Mitsuko Tanaka (University of Miami/NPF Brain Endowment Bank). Additional support was provided by the Johnnie B. Byrd Sr. Alzheimer's Disease and Research Institute, the Swiss National Science Foundation (PP00B-68859), the Verum foundation, the Bisgrove charitable donation, the NIH Neuroscience Blueprint (U24NS051872), the ENDGAME Consortium (U01HL084744), a

National Institute on Aging grant to Carl Cotman (University of California, Irvine, P50 AG23173), National Institute on Aging (KO1AG024079) and the state of Arizona.

**Please address correspondences to:** John Hardy, PhD, Reta Lila Weston Laboratories, Department of Molecular Neuroscience, Institute of Neurology, Queen Square, London WC1N3BG, England, E-mail: [j.hardy@ion.ucl.ac.uk](mailto:j.hardy@ion.ucl.ac.uk)

## References

- [1] Saunders AM, Strittmatter WJ, Schmechel D, St. George-Hyslop PH, Pericak-Vance MA, Joo SH, Rosi BL, Gusella JF, Crapper-MacLachlan DR, Alberts MJ, Hulette C, Crain B, Goldgaber D and Roses AD. Association of apolipoprotein E allele epsilon 4 with late-onset familial and sporadic Alzheimer's disease. *Neurology* 1993; 43:1467-1472.
- [2] Farrer LA, Cupples LA, Haines JL, Hyman B, Kukull WA, Mayeux R, Myers RH, Pericak-Vance MA, Risch N and van Duijn CM. Effects of age, sex and ethnicity on the association between apolipoprotein E genotype and Alzheimer disease. A meta-analysis. APOE and Alzheimer Disease Meta Analysis Consortium. *JAMA* 1997; 278:1349-1356.
- [3] Bertram L, McQueen MB, Mullin K, Blacker D and Tanzi RE. Systematic meta-analyses of Alzheimer disease genetic association studies: the AlzGene database. *Nat. Genet.* 2007; 39:17-23.
- [4] Coon KD, Myers AJ, Craig DW, Webster JA, Pearson JV, Hu-Lince D, Zismann VL, Beach T, Leung D, Bryden L, Halperin RF, Marlowe L, Kaleem M, Walker DG, Ravid R, Heward CB, Rogers J, Papassotiropoulos A, Reiman EM, Hardy J, Stephan DA. A high density whole-genome association study reveals that APOE is the major susceptibility gene for sporadic late-onset Alzheimer's disease. *J. Clin. Psychiatry* 2007; 68:613-618.
- [5] Reiman EM, Webster JA, Myers AJ, Hardy J, Dunkley T, Zismann,VL, Joshipura KD, Pearson JV, Hu-Lince D, Huentelmann MJ, Craig DW, Coon KD, Liang WS, Herbert RH, Beach T, Rohrer KC, Zhao AS, Leung D, Bryden L, Marlowe L, Kaleem M, Mastroeni D, Grover A, Heward CB, Ravid R, Rogers J, Hutton ML, Melquist S, Petersen RC, Alexander GE, Caselli RJ, Kukull W, Papassotiropoulos A, Stephan DA. GAB2 alleles modify Alzheimer's risk in APOE epsilon4 carriers. *Neuron* 2007; 54:713-720.
- [6] Hua J, Craig DW, Brun M, Webster J, Zismann V, Tembe W, Joshipura K, Huentelmann MJ, Dougherty ER, and Stephan DA. SNIPer-HD: Improved genotype calling accuracy by an expectation-maximization algorithm for high-density SNP arrays. *Bioinformatics* 2007;

- 23:57-63.
- [7] Pritchard JK, Stephens M, and Donnelly P. Inference of population structure using multilocus genotype data. *Genetics* 2000; 155:945-59.
- [8] Purcell S, Cherny SS, Sham PC. Genetic Power Calculator: design of linkage and association genetic mapping studies of complex traits. *Bioinformatics* 2003; 19:149-150.
- [9] The International HapMap Consortium. A haplotype map of the human genome. *Nature* 2005; 437:1299-1320.
- [10] Kehoe PG, Russ C, Mcllory S, Williams H, Holmans P, Holmes C, Liolitsa D, Vahidassr D, Powell J, McGleenon B, Liddell M, Plomin R, Dynan K, Williams N, Neal J, Cairns NJ, Wilcock G, Passmore P, Lovestone S, Williams J, and Owen MJ. Variation in DCP1, encoding ACE, is associated with susceptibility to Alzheimer disease. *Nat. Genet* 1999; 21:71-72.
- [11] Rigat B, Hubert C, Alhenc-Gelas F, Cambien F, Corvol P, Soubrier F. An insertion/deletion polymorphism in the angiotensin I-converting enzyme gene accounting for half the variance of serum enzyme levels. *J. Clin. Invest.* 1990; 86:1343-1346.
- [12] Sayed-Tabatabaei FA, Houwing-Duistermaat JJ, van Duijn CM, and Witteman JC. Angiotensin-converting enzyme gene polymorphism and carotid artery wall thickness: a meta-analysis. *Stroke* 2003; 34:1634-1639.
- [13] Staessen JA, Wang JG, Ginocchio G, Petrov V, Saavedra AP, Soubrier F, Vlietinck R, and Fagard R. The deletion/insertion polymorphism of the angiotensin converting enzyme gene and cardiovascular-renal risk. *J. Hypertens* 1997; 15:1579-1592.
- [14] Kivipelto M, Helkala EL, Laakso MP, Hanninen T, Hallikainen M, Alhainen K, Soininen H, Tuomilehto J, Nissinen A. Midlife vascular risk factors and Alzheimer's disease in later life: longitudinal, population based study. *BMJ* 2001; 322:1447-51.
- [15] Hofman A, Ott A, Breteler MM, Bots ML, Slooter AJ, van Harskamp F, van Duijn CN, Van Broeckhoven C, and Grobbee DE. Atherosclerosis, apolipoprotein E, and prevalence of dementia and Alzheimer's disease in the Rotterdam Study. *Lancet* 1997; 349:151-154.
- [16] Launer LJ, Ross GW, Petrovitch H, Masaki K, Foley D, White LR and Havlik RJ. Midlife blood pressure and dementia: the Honolulu-Asia aging study. *Neurobiol Aging* 2000; 21:49-55.
- [17] Roher AE, Esh C, Kokjohn TA, Kalback W, Luehrs DC, Seward JD, Sue LI, Beach TG. Circle of Willis Atherosclerosis Is a Risk Factor for Sporadic Alzheimer's Disease. *Arterioscler Thromb Vasc Biol* 2003. 23:2055:2062.
- [18] Beach TG, Wilson JR, Sue LI, Newell A, Poston M, Cisneros R, Pandya Y, Esh C, Connor DJ, Sabbagh M, Walker DG and Roher AE. Circle of Willis atherosclerosis: association with Alzheimer's disease, neuritic plaques and neurofibrillary tangles. *Acta Neuropathol (Berl)* 2007; 113:13-21
- [19] Cullen KM, Kocsi Z and Stone J. Microvascular pathology in the aging human brain: evidence that senile plaques are sites of microhaemorrhages. *Neurobiol Aging* 2006; 27:1786-96.
- [20] Cox R, Bouzekri N, Martin S, Southam L, Hugill A, Golamaully M, Cooper R, Adeyemo A, Soubrier F, Ward R, Lathrop GM, Matsuda F and Farrall M. Angiotensin-1-converting enzyme (ACE) plasma concentration is influenced by multiple ACE-linked quantitative trait nucleotides. *Hum Mol Genet* 2002; 11:2969-77.
- [21] Hardy J. Does Abeta 42 have a function related to blood homeostasis? *Neurochem Res* 2007; 32:833-5.
- [22] Fernando RL, Nettleton D, Southey BR, Dekkers JC, Rothschild MF, and Soller M. Controlling the proportion of false positives in multiple dependent tests. *Genetics* 2004; 166:611-619.
- [23] Yang Q, Cui J, Chazaro I, Cupples LA, and Demissie S. Power and type I error rate of false discovery rate approaches in genome-wide association studies. *BMC Genet.* 2005; 6:S134.
- [24] Rao DC, and Gu C. False positives and false negatives in genome scans. *Adv. Genet.* 2001; 42:487-498.
- [25] Todorov AA, and Rao DC. Trade-off between false positives and false negatives in the linkage analysis of complex traits. *Genet. Epidemiol.* 1997; 14:453-464.
- [26] Sladek R, Rocheleau G, Rung J, Dina C, Shen L, Serre D, Boutin P, Vincent D, Belisle A, Hadjadj S, Balkau B, Heude B, Charpentier G, Hudson TJ, Montpetit A, Pshezhetsky AV, Prentki M, Posner BI, Balding DJ, Meyre D, Polychronakos C and Froguel P. A genome-wide association study identifies novel risk loci for type 2 diabetes. *Nature* 2007; 445:881-5.
- [27] Diabetes Genetics Initiative of Broad Institute of Harvard and MIT, Lund University, and Novartis Institutes of BioMedical Research, Saxena R, Voight BF, Lyssenko V, Burt NP, de Bakker PI, Chen H, Roix JJ, Kathiresan S, Hirschhorn JN, Daly MJ, Hughes TE, Groop L, Altshuler D, Almgren P, Florez JC, Meyer J, Ardlie K, Bengtsson Bostrom K, Isomaa B, Lettre G, Lindblad U, Lyon HN, Melander O, Newton-Cheh C, Nilsson P, Orho-Melander M, Rastam L, Speliotes EK, Taskinen MR, Tuomi T, Guiducci C, Berglund A, Carlson J, Gianniny L, Hackett R, Hall L, Holmkvist J, Laurila E, Sjogren M, Sterner M, Surti A, Svensson M, Svensson M, Tewhey R, Blumensiel B, Parkin M, Defelice M, Barry R, Brodeur W, Camarata J, Chia N, Fava M, Gibbons J, Handsaker B, Healy C, Nguyen K, Gates C, Sougnez C, Gage D, Nizzari M, Gabriel SB, Chirn GW, Ma Q, Parikh

- H, Richardson D, Ricke D and Purcell S. Genome-wide association analysis identifies loci for type 2 diabetes and triglyceride levels. *Science* 2007; 316:1331-6.
- [28] Zeggini E, Weedon MN, Lindgren CM, Frayling TM, Elliott KS, Lango H, Timpson NJ, Perry JR, Rayner NW, Freathy RM, Barrett JC, Shields B, Morris AP, Ellard S, Groves CJ, Harries LW, Marchini JL, Owen KR, Knight B, Cardon LR, Walker M, Hitman GA, Morris AD, Doney AS; Wellcome Trust Case Control Consortium (WTCCC), Burton PR, Clayton DG, Craddock N, Deloukas P, Duncanson A, Kwiatkowski DP, Ouwehand WH, Samani NJ, Todd JA, Donnelly P, Davison D, Easton D, Evans D, Leung HT, Spencer CC, Tobin MD, Attwood AP, Boorman JP, Cant B, Everson U, Hussey JM, Jolley JD, Knight AS, Koch K, Meech E, Nutland S, Prowse CV, Stevens HE, Taylor NC, Walters GR, Walker NM, Watkins NA, Winzer T, Jones RW, McArdle WL, Ring SM, Strachan DP, Pembrey M, Breen G, St Clair D, Caesar S, Gordon-Smith K, Jones L, Fraser C, Green EK, Grozeva D, Hamshere ML, Holmans PA, Jones IR, Kirov G, Moskvina V, Nikolov I, O'donovan MC, Owen MJ, Collier DA, Elkin A, Farmer A, Williamson R, McGuffin P, Young AH, Ferrier IN, Ball SG, Balmforth AJ, Barrett JH, Bishop DT, Iles MM, Maqbool A, Yuldashova N, Hall AS, Braund PS, Dixon RJ, Mangino M, Stevens S, Thompson JR, Bredin F, Tremelling M, Parkes M, Drummond H, Lees CW, Nimmo ER, Satsangi J, Fisher SA, Forbes A, Lewis CM, Onnie CM, Prescott NJ, Sanderson J, Mathew CG, Barbour J, Mohiuddin MK, Todhunter CE, Mansfield JC, Ahmad T, Cummings FR, Jewell DP, Webster J, Brown MJ, Lathrop GM, Connell J, Dominiczak A, Braga Marciano CA, Burke B, Dobson R, Gungadoo J, Lee KL, Munroe PB, Newhouse SJ, Onipinla A, Wallace C, Xue M, Caulfield M, Farrall M, Barton A, Bruce IN, Donovan H, Eyre S, Gilbert PD, Hider SL, Hinks AM, John SL, Potter C, Silman AJ, Symmons DP, Thomson W, Worthington J, Dunger DB, Widmer B, Newport M, Sirugo G, Lyons E, Vannberg F, Hill AV, Bradbury LA, Farrar C, Pointon JJ, Wordsworth P, Brown MA, Franklyn JA, Heward JM, Simmonds MJ, Gough SC, Seal S, Stratton MR, Rahman N, Ban M, Goris A, Sawcer SJ, Compston A, Conway D, Jallow M, Rockett KA, Bumpstead SJ, Chaney A, Downes K, Ghorri MJ, Gwilliam R, Hunt SE, Inouye M, Keniry A, King E, McGinnis R, Potter S, Ravindrarajah R, Whittaker P, Widdon C, Withers D, Cardin NJ, Ferreira T, Pereira-Gale J, Hallgrimsdottir IB, Howie BN, Su Z, Teo YY, Vukcevic D, Bentley D, Compston A, Ouwehand NJ, Samani MR, Isaacs JD, Morgan AW, Wilson GD, Arder-Jones A, Berg J, Brady A, Bradshaw N, Brewer C, Brice G, Bullman B, Campbell J, Castle B, Cetnarskyj R, Chapman C, Chu C, Coates N, Cole T, Davidson R, Donaldson A, Dorkins H, Douglas F, Eccles D, Eeles R, Elmslie F, Evans DG, Goff S, Goodman S, Goudie D, Gray J, Greenhalgh L, Gregory H, Hodgson SV, Homfray T, Houlston RS, Izatt L, Jackson L, Jeffers L, Johnson-Roffey V, Kavalier F, Kirk C, Laloo F, Langman C, Locke I, Longmuir M, Mackay J, Magee A, Mansour S, Miedzybrodzka Z, Miller J, Morrison P, Murday V, Paterson J, Pichert G, Porteous M, Rahman N, Rogers M, Rowe S, Shanley S, Saggat A, Scott G, Side L, Snadden L, Steel M, Thomas M, Thomas S, McCarthy MI and Hattersley AT. Replication of genome-wide association signals in UK samples reveals risk loci for type 2 diabetes. *Science* 2007; 316:1336-41.
- [29] Scott LJ, Mohlke KL, Bonnycastle LL, Willer CJ, Li Y, Duren WL, Erdos MR, Stringham HM, Chines PS, Jackson AU, Prokunina-Olsson L, Ding CJ, Swift AJ, Narisu N, Hu T, Pruim R, Xiao R, Li XY, Conneely KN, Riebow NL, Sprau AG, Tong M, White PP, Hetrick KN, Barnhart MW, Bark CW, Goldstein JL, Watkins L, Xiang F, Saramies J, Buchanan TA, Watanabe RM, Valle TT, Kinnunen L, Abecasis GR, Pugh EW, Doherty KF, Bergman RN, Tuomilehto J, Collins FS and Boehnke M. A genome-wide association study of type 2 diabetes in Finns detects multiple susceptibility variants. *Science* 2007; 316:1341-5.
- [30] Blomqvist ME-L, Reynolds C, Katzov H, Feuk L, Andreasen N, Bogdanovic N, Blennow K, Brookes AJ, and Prince JA. Towards compendia of negative genetic association studies: an example for Alzheimer disease. *Hum. Genet.* 2006; 119:29-37.
- [31] Goring HH, Terwilliger JD, and Blangero J. Large upward bias in estimation of locus-specific effects from genomewide scans. *Am. J. Hum. Genet.* 2001; 69:1357-1369.
- [32] Lohmueller KE, Pearce CL, Pike M, Lander ES, and Hirschhorn JN. Meta-analysis of genetic association studies supports a contribution of common variants to susceptibility to common disease. *Nat. Genet.* 2003; 33:177-182.
- [33] Kavvoura FK, McQueen MB, Khoury MJ, Tanzi RE, Bertram L, Ioannidis JP. Evaluation of the potential excess of statistically significant findings in published genetic association studies: application to Alzheimer's disease. *Am J Epidemiol.* 2008 Oct 15;168(8):855-65. Epub 2008 Sep 8.
- [34] Thornton-Wells TA, Moore JH, Martin ER, Pericak-Vance MA, Haines JL. Confronting complexity in late-onset Alzheimer disease: application of two-stage analysis approach addressing heterogeneity and epistasis. *Genet Epidemiol.* 2008 Apr;32(3):187-203
- [35] Lii H, Wetten S, Li L, St Jean PL, Upmanyu R, Surh L, Hosford D, Barnes MR, Briley JD, Borrie M, Coletta N, Delisle R, Dhalla D, Ehm MG, Feldman HH, Fornazzari L, Gauthier S, Goodgame N, Guzman D, Hammond S, Hollingworth P, Hsiung GY, Johnson J, Kelly DD, Keren R, Kertesz A, King KS, Lovestone S, Loy-

## ACE as AD Risk Factor

English I, Matthews PM, Owen MJ, Plumpton M, Pryse-Phillips W, Prinjha RK, Richardson JC, Saunders A, Slater AJ, St George-Hyslop PH, Stinnett SW, Swartz JE, Taylor RL, Wherrett J, Williams J, Yarnall DP, Gibson RA, Irizarry MC,

Middleton LT, Roses AD. Candidate single-nucleotide polymorphisms from a genomewide association study of Alzheimer disease. *Arch Neurol.* 2008 Jan;65(1):45-53.